# Overall view of the ACORD project

Gabriel G. Bès, Thierry Guillotin

G. G. Bès T. Guillotin (Eds.)

# A Natural Language and Graphics Interface

## Results and Perspectives from the ACORD Project

Springer-Verlag
Berlin Heidelberg New York London Paris
Tokyo Hong Kong Barcelona Budapest

CHAPTER 1

# Overall View of the ACORD Project

Thierry Guillotin, LdM
Gabriel G. Bès, CLF

## 1 General Architecture

The ACORD system combines Natural Language and Graphic techniques for the construction and interrogation of knowledge bases in French, German and English. The selected domain of application is a knowledge base concerning transport services. The ACORD demonstrator integrates all the components of the project. It is the coherent result of the theoretical and technical work performed in the project. Each of its components can perform separately.

The major components of the system, represented in figure 1, are the following :

- parsers for the three languages
- a graphics component
- a knowledge base and a theorem prover
- text generators for the three languages
- a dialogue manager including a resolver and a planner

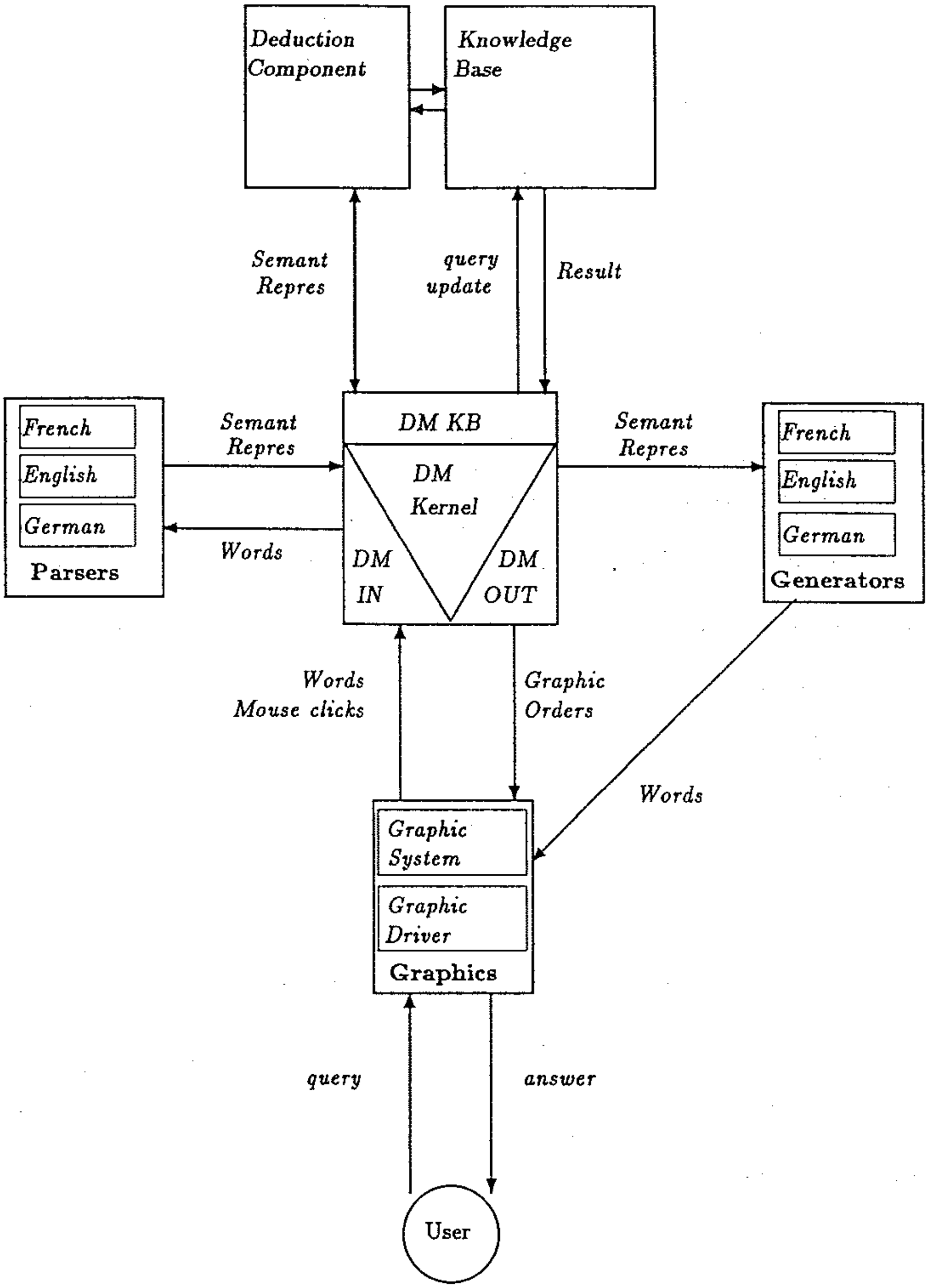
### 1.1 The Parsers

Two different syntactic frameworks and parsing tools have been investigated in the general framework of unification grammar. The English and French grammars are designed using UCG (Unification Categorial Grammars). The German grammar is designed using LFG (Lexical Functional Grammar). The three parsers deliver their output in a common KRL language, thus rendering the multilingual aspect of the system transparent to the other components. This language, called InL (INdexed Language), has been developed for ACORD as a derivative from DRT (Discourse Representation Theory).

### 1.2 The Graphics Component

The aim of the graphics system is to allow the user to mix graphics and text in a flexible way; the system understands the semantics of such mixed expressions. Moreover, the system can also combine graphics and text when displaying results. Such an integration of graphics and natural language requires a common means of communication between the two components. The language used is InL and its derivatives which are analysed both by the parsers and the graphics system and produced both by the generators and the graphics system. The functionalities of this are :

- deixis pointing : the graphics component displays a map of the current state of the knowledge base. The user can refer to any object on this map by using deictics such as *this, that, here* and




*DM : Dialogue Manager*

Figure 1: ACORD System Architecture

*there* in a sentence and he can specify the referred object by clicking on it with the mouse.

• highlighting : the graphics component can also collaborate to query answering by highlighting an object on the map.

• direct manipulation : the user can move objects on the map and also update the values of the charts associated to truck loads or city depots.

### 1.3 The Knowledge Base and the Theorem Prover

The theorem prover and the knowledge base are responsible for storing and retrieving the information entered by the user. The theorem prover stands as a filter to the knowledge base, retaining general assertions, providing help with complex queries and also formating the input to the knowledge base. The knowledge base receives as input a DRS and records the information in this formalism (DRS is a syntactic variant of InL). It performs a consistency check on each item of asserted information verifying, in this way, the stored knowledge. To answer the user's queries the knowledge base uses a semantic network to retrieve generic knowledge and modify the initial DRS from query, enabling the inference process from queries to known facts.

### 1.4 The Generators

The system provides short answers in French, German or English depending on the language used in the query. The first version of the generator was focused on generation viewed as inverse parsing. The second version is based on a planning component which enables the construction of appropriate referential terms.

### 1.5 The Dialogue Manager

The Dialogue Manager supervises the whole system. It controls communications between the other subcomponents of ACORD. The interfaces can be briefly described in the following table.

| Modules | Input from the DM | Output to the DM |
|---|---|---|
| Parsers | List of words | InL with resolution information |
| Generators | SynInl | List of words |
| Graphics system | DRS updates | DRS for KB updates |
| Theorem prover | DRS | Results from KB retrieval |

Besides scheduling the transmission of information from one component to the other and processing the output of one component to fit the requierements of the following components, it is also in charge of the resolver and planner processes. The first one handles anaphora resolution for the three languages, the latter specifies the input to the generators.

## 2 The Contributions

With the above overall picture of the project in mind, each of the following chapters can be read independently of the others. In some cases (see specialy chapter 8 §5), the relations of a particular module with the others is explained in more detail than previously, but with minimum overlapping.

The three following chapters develop specific issues of the grammar models used in the ACORD project. In chapter 2, Walter Kasper defines a solution to the construction of semantic representations in the LFG pattern ; he proposes some modifications to LFG in order to obtain results more effectively. Chapters 3 and 4 are devoted to UCG grammars. In chapter 3, Karine Baschung, Gabriel G. Bès and Thierry Guillotin present two solutions to the adaptation of UCG to French specific data. Both of them introduce interesting extensions of the original model. The first one introduces (§4.1.5) co-relational constraints, which basically allow the expression of disjunction over sets of variables. The second one considers argumental valencies as being organized in sets rather than in lists (§4.2). In chapter 4, Pierre-François Jurie and Gabriel G. Bès suggest that linguistic observations can be captured by an axiomatic system ; models satisfying the axioms can be compared with sentences specified by the grammars in order to control their descriptive adequacy. They show how this can be formerly performed with a UCG type grammar.

Each module of the ACORD system can perform separately, but the whole system is highly structured and integrated. The basic challenge was to define for some linguistic processes the adequate trade-off between accounting for the particularities of each language and the specification of cross linguistic generalisations, inasmuch as the three linguistic modules obey common features and/or are required to access the same sources of knowledge. The next two chapters illustrate significant results of the ACORD project in this area.

In chapter 5, Walter Kasper, Marc Moens and Henk Zeevat present in a detailed manner the interesting solutions of the ACORD project with respect to anaphora resolution in three languages. They explain carefully what is language dependent and what is system and domain dependent, and how they managed to tackle modularly the whole problem.

Dieter Kohl, Claire Gardent, Agnès Plainfossé, Mike Reape and Stefan Momma explain carefully in chapter 6 the ACORD solution for generation. This defines a common cross-linguistic part and specific generators for the three languages. The former incorporates a planner and is in charge of adding syntactically motivated structure, common to the three languages, to InL representations, which become thus SynInl ones.

Another challenge of the ACORD project was the integration of graphics and natural language semantics. In chapter 7, John Lee explains the underlying conceptual difficulties of graphical communication, indicates how the integration was obtained within the limits of the system and presents the ACORD graphic architecture.

In the final chapter -chapter 8- Gerhard Heyer, Friedrich Dudda and Peter Ovenhausen, discuss basic issues of the knowledge base module. The crucial challenge here was to define a knowledge representation formalism independent of specific languages and specific mode, and powerful enough to account both for queries and updates, and for factual and conceptual knowledge.

## 3 The Teams

The ACORD system is the result of five years work by eight European teams, industrial and academic partners, with the involvement of more than fifty people. The participating teams were the following.[1]

**LdM** — **Alcatel Alsthom Recherche**
(formerly Laboratoires de Marcoussis)
Route de Nozay, 91460 Marcoussis, France

| | |
|---|---|
| ECCS | University of Edinburgh, Centre for Cognitive Science<br>2 Buccleuch Place, Edinburgh EH8 9LW, Scotland, UK |
| EdCAAD | University of Edinburgh, Department of Architecture<br>20 Chambers Street, Edinburgh EH1 1JZ, Scotland, UK |
| CLF | Université Blaise Pascal (Clermont Ferrand II)<br>Formation Doctorale Linguistique et Informatique<br>34 Avenue Carnot, 63037 Clermont Ferrand Cedex, France |

**TA** — TA Triumph-Adler
Hundingstr. 11 bd, 8500 Nuremberg, Germany

| | |
|---|---|
| FhG | Fraunhofer Institut fur Arbeitswirtschaft und Organisation<br>Holzgartenstr. 17, 7000 Stuttgart 1, Germany |

**BULL** — Bull-Cediag
68 route de Versailles, 78430 Louveciennes, France

| | |
|---|---|
| IMS | Universitat Stuttgart, Institut fur Maschinelle Sprachverarbeitung-Cl<br>Azenberstr. 17, 7000 Stuttgart 1 , Germany |

It is hoped that the following chapters will offer to the reader something more than interesting results and perspectives from the ACORD project. An attentive reading will also permit the discovery of the internal positive dynamics of the whole project, which become more and more integrated around the ACORD demonstrator. Final improvements were obtained due to a very collaborative way of working.

[1] Some of the contributors to this volume are now (end 1991) working at different sites. They are Claire Gardent (Dept. of Philosophy, Faculteit der Wijsbegeerte, Heidelberglaan, 2 Postbus 80103, 3508 TC Utrecht), Thierry Guillotin (Centre Scientifique d'IBM France, 3-5 Place Vendôme, F 75021 Paris Cedex 01), Mike Reape (Departement of Computer Science, O'Reilly Institute, Trinity College, Dublin 2, Eire)